1. Muestreo aleatorio simple

1. Muestreo aleatorio simple.

1.0 Definiciones y conceptos básicos.
1.1 Selección de una muestra aleatoria simple. Números aleatorios. Rutas aleatorias.
1.2 Muestreo aleatorio simple en poblaciones infinitas.
  1.2.1 Media, varianza y proporción muestrales: Propiedades. Error de estimación
  1.2.2 Estimación puntual. Intervalos de confianza. Contrastes de hipótesis.
  1.2.3 Determinación del tamaño muestral.
1.3 Muestreo aleatorio simple en poblaciones finitas.
  1.3.1 Estimación de la media, proporción y total poblacionales.
  1.3.2 Determinación del tamaño muestral.

1.0 Definiciones y conceptos básicos

Muestreo aleatorio simple: Todas las muestras de un determinado tamaño tienen la misma probabilidad de ser seleccionadas. Cada individuo de la población tiene la misma probabilidad de ser seleccionado para la muestra.

Muestreo aleatorio estratificado: Se divide a la población en grupos, denominados estratos, y se seleccionar una muestra aleatoria simple de cada estrato.

Muestreo por conglomerados: Se divide a la población en grupos, denominados conglomerados, y seleccionar una muestra aleatoria simple de conglomerados.

Muestreo sistemático: Consiste en seleccionar un elemento al comienzo de una lista de la población y luego se selecciona cada un número fijo de posiciones el resto de elementos.

Muestreo con probabilidades desiguales: Las probabilidades de selección de la muestra y de los distintos individuos son desiguales.

1. Muestreo aleatorio simple

1.0 Definiciones y conceptos básicos

Errores de no muestreo

Sesgo de selección. Ocurre cuando alguna parte de la población objetivo no está en la población muestreada.

Sesgo de medición. Ocurre cuando por diversos motivos los datos que obtenemos no son exactos o verdaderos.

No respuesta. La no respuesta de individuos seleccionados para la muestra puede causar sesgo en los datos muestrales.

1. Muestreo aleatorio simple

1.0 Definiciones y conceptos básicos

Razones para el uso del muestreo

Evitar la destrucción de la población.

En algunos casos, una unidad de observación debe ser destruida para ser observada. En ese caso, un censo destruiría toda la población.

Rapidez.

Los datos se pueden reunir más rápido, de modo que las estimaciones se pueden publicar de una manera programada

Economía y precisión.

El muestreo puede proporcionar información fiable con costes mucho menores que los de un censo.

1. Muestreo aleatorio simple

1.1 Selección de una muestra aleatoria simple. Números aleatorios. Rutas aleatorias.

- Cada muestra posible de tamaño n tiene la misma probabilidad de ser seleccionada.
- **En la práctica, la anterior condición se traduce en que cada elemento tenga la misma probabilidad de pertenecer a la muestra.**
- La selección de cada elemento de la muestra se hace sobre la base de un sorteo completamente aleatorio.
- Opciones: tablas de números aleatorios o generación de números aleatorios con programas de ordenador.

1107 1032 2596 4562 7598 1546 2596 5412 8569 2563 4587 2596 5641 5866 5844

2687 1596 3589 6578 1452 2365 7899 4122 1477 8836 3696 2587 6985 5632 5896

....

- Ordenar o numerar la población.
- Rutas aleatorias (guías telefónicas)

IMPORTANTE: EL NÚMERO TOTAL DE ELEMENTOS QUE FORMAN UNA MUESTRA TIENE MENOS IMPORTANCIA QUE EL PRINCIPIO DE SELECCIÓN ALEATORIA

1. Muestreo aleatorio simple

1.2 Muestreo aleatorio simple en poblaciones infinitas

1.2.1. Media, varianza y proporción muestrales: Propiedades. Error de estimación.

Población Y (μ,σ^2). Muestra aleatoria simple $Y_1,\ldots,Y_n$ **(i.i.d.)**

$$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i \qquad E(\bar{y})=\mu \qquad V(\bar{y})=\frac{\sigma^2}{n} \qquad S^2=S_{n-1}^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2 \qquad E(S^2)=\sigma^2$$

$$\hat{V}(\bar{y})=\frac{S^2}{n}$$

Cuando las variables Y, $Y_1,\ldots,Y_n$ son dicotómicas $\mu=p$, $\sigma^2=pq$

$$\hat{p}=\frac{1}{n}\sum_{i=1}^{n}y_i,\quad y_i=0,1 \qquad E(\hat{p})=p \qquad V(\hat{p})=\frac{\sigma^2}{n}=\frac{pq}{n} \qquad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2=\frac{n}{n-1}\hat{p}\hat{q} \qquad E(S^2)=pq$$

$$\hat{V}(\hat{p})=\frac{\hat{p}\hat{q}}{n-1}$$

1. Muestreo aleatorio simple

1.2 Muestreo aleatorio simple en poblaciones infinitas

1.2.1. Media, varianza y proporción muestrales: Propiedades. Error de estimación.

$Y \to N(\mu,\sigma^2) \quad \sigma^2 \text{ conocida} \qquad \dfrac{\bar{y}-\mu}{\frac{\sigma}{\sqrt{n}}} \to N(0,1)$

$Y \to N(\mu,\sigma^2) \quad \sigma^2 \text{ desconocida} \qquad \dfrac{\bar{y}-\mu}{\frac{S}{\sqrt{n}}} \to t_{n-1} \approx N(0,1) \quad (para \;\; n>30)$

$Y \to cualquier\ ley \quad (n\to\infty) \qquad \dfrac{\bar{y}-\mu}{\frac{\sigma}{\sqrt{n}}} \approx \dfrac{\bar{y}-\mu}{\frac{S}{\sqrt{n}}} \to N(0,1) \quad (para \;\; n>30)$

$Y \to B(1,p) \quad \mu = p \quad \bar{y} = \hat{p} \qquad \dfrac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} \approx \dfrac{\hat{p}-p}{\sqrt{\frac{n}{n-1}\frac{\hat{p}\hat{q}}{n}}} = \dfrac{\hat{p}-p}{\sqrt{\frac{\hat{p}\hat{q}}{n-1}}} \to N(0,1)$

7

1. Muestreo aleatorio simple

1.2 Muestreo aleatorio simple en poblaciones infinitas

1.2.1. Media, varianza y proporción muestrales: Propiedades. Error de estimación.

$$P\left[-2\frac{\sigma}{\sqrt{n}} \le \bar{y}-\mu \le 2\frac{\sigma}{\sqrt{n}}\right] = 0{,}95 \quad \Rightarrow \quad P\left[\left|\bar{y}-\mu\right| \le 2\frac{\sigma}{\sqrt{n}}\right] = 0{,}95$$

$N(0,1)$

$$P\left[-Z_{\alpha/2} \le \frac{\bar{y}-\mu}{\frac{\sigma}{\sqrt{n}}} \le Z_{\alpha/2}\right] = 1-\alpha \qquad P\left[-1{,}96 \le \frac{\bar{y}-\mu}{\frac{\sigma}{\sqrt{n}}} \le 1{,}96\right] = 0{,}95$$

$\boxed{1{,}96 \approx 2}$

$$P\left[\bar{y}-2\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{y}+2\frac{\sigma}{\sqrt{n}}\right] = 0{,}95 \qquad \left(\bar{y}-2\frac{\sigma}{\sqrt{n}},\ \bar{y}+2\frac{\sigma}{\sqrt{n}}\right)$$

8

1. Muestreo aleatorio simple

1.2 Muestreo aleatorio simple en poblaciones infinitas

1.2.1. Media, varianza y proporción muestrales: Propiedades. Error de estimación.

$$P\left[\left|\bar{y}-\mu\right|\leq 2\frac{\sigma}{\sqrt{n}}\right]=0,95$$

Desigualdad de Tchebychev

$$E(X)=\mu \qquad V(X)=\sigma^2 \qquad \Rightarrow \qquad P\left[\left|X-\mu\right|\leq k\sigma\right]\geq 1-\frac{1}{k^2}$$

$$E(\bar{y})=\mu \quad V(\bar{y})=\frac{\sigma^2}{n} \quad k=2 \qquad \Rightarrow \qquad P\left[\left|\bar{y}-\mu\right|\leq 2\frac{\sigma}{\sqrt{n}}\right]\geq 1-\frac{1}{4}=0,75$$

9

9

1. Muestreo aleatorio simple

1.2 Muestreo aleatorio simple en poblaciones infinitas

1.2.2 Estimación puntual. Intervalos de confianza. Contrastes de hipótesis.

ESTIMACIÓN PUNTUAL $\quad \bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i \qquad 2\sqrt{\hat{V}(\bar{y})}=2\frac{S}{\sqrt{n}}$

INTERVALO DE CONFIANZA $\quad \left(\bar{y}-2\frac{S}{\sqrt{n}}\,,\,\bar{y}+2\frac{S}{\sqrt{n}}\right)$

CONTRASTE DE HIPÓTESIS

- $se\ acepta\ H_0:\mu=\mu_0 \quad si \quad \mu_0\in\left(\bar{y}-2\frac{S}{\sqrt{n}},\bar{y}+2\frac{S}{\sqrt{n}}\right)$
- $se\ rechaza\ H_0 \quad si \quad \mu_0\notin\left(\bar{y}-2\frac{S}{\sqrt{n}},\bar{y}+2\frac{S}{\sqrt{n}}\right)$

10

10

1. Muestreo aleatorio simple

1.2 Muestreo aleatorio simple en poblaciones infinitas

1.2.3 Determinación del tamaño muestral.

$$2\sqrt{V(\overline{y})} = 2\sqrt{\frac{\sigma^2}{n}} = B \qquad 4\frac{\sigma^2}{n} = B^2 \Rightarrow n = \frac{\sigma^2}{\frac{B^2}{4}} = \frac{\sigma^2}{D} \quad , \quad D = \frac{B^2}{4}$$

$$\sigma^2 = S^2 \qquad \sigma \cong \frac{R}{4} \Leftrightarrow \sigma^2 \cong \frac{R^2}{16}$$

$$2\sqrt{V(\hat{p})} = 2\sqrt{\frac{pq}{n}} = B \qquad n = \frac{pq}{\frac{B^2}{4}} = \frac{pq}{D} \quad , \quad D = \frac{B^2}{4}$$

$$p = \hat{p} \qquad p = q = \frac{1}{2}$$

11

1. Muestreo aleatorio simple

1.3 Muestreo aleatorio simple en poblaciones finitas

1.3.1 Estimación de la media, proporción y total poblacionales.

$$\hat{\mu} = \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad E(\overline{y}) = \mu \qquad V(\overline{y}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) \qquad E(S^2) = \frac{N}{N-1}\sigma^2 \qquad E\left(\frac{N-1}{N}S^2\right) = \sigma^2$$

$$\hat{V}(\overline{y}) = \frac{S^2}{n}\left(\frac{N-n}{N}\right)$$

$$\left(\frac{N-n}{N}\right) \qquad \left(\frac{N-n}{N}\right) \geq 0{,}95 \Leftrightarrow n \leq \frac{1}{20}N = 5\%N$$

$$\hat{\tau} = N\overline{y} = \frac{N}{n}\sum_{i=1}^{n} y_i \qquad \hat{V}(\hat{\tau}) = \hat{V}(N\overline{y}) = N^2\hat{V}(\overline{y}) = N^2\frac{S^2}{n}\frac{N-n}{N} = N(N-n)\frac{S^2}{n}$$

12

1. Muestreo aleatorio simple

1.3 Muestreo aleatorio simple en poblaciones finitas

1.3.1 Estimación de la media, proporción y total poblacionales.

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad y_i = 0,1 \qquad S^2 = \frac{n\hat{p}\hat{q}}{n-1} \qquad \hat{V}(\hat{p}) = \frac{S^2}{n}\frac{N-n}{N} = \frac{\hat{p}\hat{q}}{n-1}\frac{N-n}{N}$$

$$\hat{\tau} = N\hat{p} \qquad \hat{V}(\hat{\tau}) = \hat{V}(N\hat{p}) = N^2\hat{V}(\hat{p}) = N(N-n)\frac{\hat{p}\hat{q}}{n-1}$$

13

13

1. Muestreo aleatorio simple

1.3 Muestreo aleatorio simple en poblaciones finitas

1.3.2. Determinación del tamaño muestral.

$$2\sqrt{\frac{\sigma^2}{n}\frac{N-n}{N-1}} = B \Rightarrow n = \frac{N\sigma^2}{(N-1)D+\sigma^2} \qquad D = \frac{B^2}{4} \ (media) \qquad D = \frac{B^2}{4N^2} \ (total)$$

$$\sigma^2 = S^2$$

$$\sigma \cong \frac{R}{4} \Leftrightarrow \sigma^2 \cong \frac{R^2}{16}$$

$$n = \frac{Npq}{(N-1)D+pq} \qquad D = \frac{B^2}{4} \ (proporcion) \qquad D = \frac{B^2}{4N^2} \ (total)$$

$$p = \hat{p}$$

$$p = q = \frac{1}{2}$$

14

14

1. Muestreo aleatorio simple

1.3 Muestreo aleatorio simple en poblaciones finitas

Ejemplo donde podemos observar todas las propiedades de los estimadores estudiadas en el muestreo aleatorio simple sobre poblaciones finitas.

	Población finita	Proporción poblacional, p (de negras)	Media poblacional, μ	Varianza poblacional, σ^2
①②③ ⓫⓬⓭	N=6	$p=\frac{3}{6}=0,50$	$\mu=\frac{1+2+3+11+12+13}{6}=7$	25,6667

1. Muestreo aleatorio simple

1.3 Muestreo aleatorio simple en poblaciones finitas

	(no escojo)	**MUESTRA (n=4)**	Proporción muestral, $\hat{p}$ (de negras)	Media muestral, $\bar{y}$	S^2_{n-1}
1	①②	③⓫⓬⓭	$\frac{3}{4}=0,75$	$\frac{3+11+12+13}{4}=9,75$	20,92
2	①③	②⓫⓬⓭	$\frac{3}{4}=0,75$	$\frac{2+11+12+13}{4}=9,5$	25,67
3	①⓫	②③⓬⓭	$\frac{2}{4}=0,50$	$\frac{2+3+12+13}{4}=7,5$	33,67
4	①⓬	②③⓫⓭	$\frac{2}{4}=0,50$	$\frac{2+3+11+13}{4}=7,25$	30,92
5	①⓭	②③⓫⓬	$\frac{2}{4}=0,50$	$\frac{2+3+11+12}{4}=7$	27,33
6	②③	①⓫⓬⓭	$\frac{3}{4}=0,75$	$\frac{1+11+12+13}{4}=9,25$	30,92
7	②⓫	①③⓬⓭	$\frac{2}{4}=0,50$	$\frac{1+3+12+13}{4}=7,25$	37,58
8	②⓬	①③⓫⓭	$\frac{2}{4}=0,50$	$\frac{1+3+11+13}{4}=7$	34,67
9	②⓭	①③⓫⓬	$\frac{2}{4}=0,50$	$\frac{1+3+11+12}{4}=6,75$	30,92
10	③⓫	①②⓬⓭	$\frac{2}{4}=0,50$	$\frac{1+2+12+13}{4}=7$	40,67
11	③⓬	①②⓫⓭	$\frac{2}{4}=0,50$	$\frac{1+2+11+13}{4}=6,75$	37,58
12	③⓭	①②⓫⓬	$\frac{2}{4}=0,50$	$\frac{1+2+11+12}{4}=6,5$	33,67
13	⓫⓬	①②③⓭	$\frac{1}{4}=0,25$	$\frac{1+2+3+13}{4}=4,75$	30,92
14	⓫⓭	①②③⓬	$\frac{1}{4}=0,25$	$\frac{1+2+3+12}{4}=4,5$	25,67
15	⓬⓭	①②③⓫	$\frac{1}{4}=0,25$	$\frac{1+2+3+11}{4}=4,25$	20,92
		TOTAL:	7,50	105	462

1. Muestreo aleatorio simple

1.3 Muestreo aleatorio simple en poblaciones finitas

MUESTRAS (*n*=4)	Proporción muestral, $\hat{p}$ (de negras)	Media muestral, $\bar{y}$	S_{n-1}^2
TOTAL:	7,50	105	462
MEDIA:	$E\left[\hat{p}\right]=\frac{7,5}{15}=0,50$	$E\left[\bar{y}\right]=\frac{105}{15}=7$	30,8
VARIANZA:	$V\left[\hat{p}\right]=0,025$	$V\left[\bar{y}\right]=2,5667$	

17

17

1. Muestreo aleatorio simple

1.3 Muestreo aleatorio simple en poblaciones finitas

	(no escojo)	**MUESTRAS (*n*=5)**	Proporción muestral, $\hat{p}$ (de negras)	Media muestral, $\bar{y}$	S_{n-1}^2
1	①	②③⓫⓬⓭	$\frac{3}{5}=0,60$	$\frac{2+3+11+12+13}{5}=8,2$	27,70
2	②	①③⓫⓬⓭	$\frac{3}{5}=0,60$	$\frac{1+3+11+12+13}{5}=8$	31,00
3	③	①②⓫⓬⓭	$\frac{3}{5}=0,60$	$\frac{1+2+11+12+13}{5}=7,8$	33,70
4	⓫	①②③⓬⓭	$\frac{2}{5}=0,40$	$\frac{1+2+3+12+13}{5}=6,2$	33,70
5	⓬	①②③⓫⓭	$\frac{2}{5}=0,40$	$\frac{1+2+3+11+13}{5}=6$	31,00
6	⓭	①②③⓫⓬	$\frac{2}{5}=0,40$	$\frac{1+2+3+11+12}{5}=5,8$	27,70
		TOTAL:	3	42	184,8
		MEDIA:	$E\left[\hat{p}\right]=\frac{3}{6}=0,50$	$E\left[\bar{y}\right]=\frac{42}{6}=7$	30,8
		VARIANZA:	$V\left[\hat{p}\right]=0,01$	$V\left[\bar{y}\right]=1,0267$	

18

18